Project Title: Phoebus
PI: Martin Swany
Institution: University of Delaware

Project Summary:

The goal is the Phoebus is to explore next-generation network environments for high-performance distributed computing. Both the core network protocols and the way in which networks are used have been relatively static for a number of years. The emergence of environments such as the DOE UltraScienceNet, driven by the ever-increasing demands on the network by DOE science applications, indicate that the community must continue to investigate new network architectures to ensure that the needs of applications are met. Serving that interest, we are investigating a novel set of protocols and middleware to help build this bridge to the future of high-impact science.

The core of this work is the development and investigation of new network services, exposed via a "session" layer. As we have indicated in other work, releasing applications from a direct dependence on an end to end transport layer connection offers many opportunities for functional and performance improvements. Essentially the historic binding of transport to network layer stems from the decision to superficially hide all differences in various network components and take a "one size fits all" approach. In the modern world, we must deal with heterogeneity of network technology and administrative domains and policies. The essence of the functional improvements can be described as providing a framework for explicit negotiation as a part of connection establishment that can adapt an end to end communication based on the various networks it traverses. Instead of a simple, end-to-end connection, end-systems would have a session view that links together the services in each network, and allows us to deal with the heterogeneity that exists between networks.

Another cornerstone of this investigation is that it is targeted for implementation on network processors. These systems permit complex packet handling operations at line rate speeds. Network processors have gotten significantly more powerful over the generations, currently able to operate at 10Gb/sec, and there is no reason to think that this trend would not continue. Based on their current viability and their future capabilities, these systems seem like perfect matches for implementing the ideas of the Phoebus project in an easy to deploy setup, and thus, are a major focus of the middleware to be produced by the project. The project has a preliminary implementation on the Intel IXP 2400 network processor, and we are investigating techniques that are viable in this environment for transport protocol and control signal modification.

Research Activities:

- Session protocol design and implementation – The Generic Session Protocol has been revised and Version 1 is nearing completion. This protocol is extensible via the use of options, but it currently explicitly targets authentication and authorization, transport and traffic engineering transformation points, as well as the negotiation based on the intersection of current network conditions and policy.

- Preliminary implementation experience on Intel IXP2400 NP platform – We have a functioning Intel IXP2400 development environment and a basic working prototype. The IXP system that we have features 3 Gbit Ethernet interfaces only but it contains the same micro-architecture as the newer IXP 2800, which can support 10Gbit Ethernet as well as SONET OC-192.

- GSI authentication implementation and protocol support – Protocol and implementation support for GSI authentication in the session framework. This functionality essentially allows an end system to dynamically open forwarding channels in the network.

- Link-state routing protocol to exchange flow information inside an administrative domain to enable policy evaluation – This protocol is designed to address the case where networks peer in multiple locations and when the admission of new flows should be sensitive to the current traffic patterns between domains and their associated policy.

- SAML-based models for policy – We have experimented with using the Security Assertion Markup Language to allow indirection in the way in which resources are authorized. Thus rather than have network access tied to an IP address or user, it can be allowed based on site-based certificates.

- Transport protocol translation investigation on the IXP – We have begun work on the design and implementation and of line-speed TCP termination for multi-transport connections. This would allow a network device to do transport-layer transcoding within a session. This allows from a single transport connection to a series of shorter, better-performing TCP connections, or utilization of a rate-based protocol in parts of the network where flows are isolated.

- Investigation and evaluation of signaling an traffic engineering translation points – We have begun to investigate support for RSVP-TE and the ability to evaluate the addition of Explict Route options to the Path messages. In addition, we have begun to work with models where DSCPs and MPLS labels can be translated at administrative boundaries.

DOE Collaborators: Micah Beck. Les Cottrell, Harvey Newman, Joe Metzger, Don Petravick, Nagi Rao, Bill Wing

Impact to specific DOE science applications: While the problem of Tier 0 to Teir 1 sites for the ATLAS and CMS experiments has been largely solved, the question of the how the Tier 2 sites will get access to experiment data remains open. The Tier 2 centers in the US will almost certainly connect over various networks including a mixture of Internet2's Abilene and peering with the Labs over MANs. This technology stands to address many emerging issues.

Synergy developed with DOE application developers to facilitate technology transfers: Plans have been discussed with both ORNL and FNAL for technology transfer. Technology transfer has also been discussed with CERN in Geneva.